MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
Zaporizhzhia Polytechnic National University

WORKBOOK ON HIGHER MATHEMATICS (4th module)

for students majoring in
141 Power Engineering, Electrical Engineering
and Electrical Mechanics

2021

Workbook on higher mathematics (4th module) for students majoring in 141 Power Engineering, Electrical Engineering and Electrical Mechanics / Compiler: Snizhko Nataliia. – Zaporizhzhia: Zaporizhzhia Polytechnic National University, 2021. – 46 p.

Compiler: Snizhko Nataliia,
candidate of science (physics and mathematics),
associate professor

Executive editor: Snizhko Nataliia,
candidate of science (physics and mathematics),
associate professor

Reviewer: Antonenko Nina.
candidate of science (physics and mathematics),
associate professor

Major expert: Antonov Mykola,
candidate of science (engineering),
associate professor

English language adviser: Khavkina Olena,
candidate of science (philology),
associate professor

Approved
by the Department of Higher Mathematics
(Minutes № 3 of 21.10.2020)

Recommended
by the Scientific and Methodological Council
of the Faculty of Mechanical Engineering
(Minutes № 4 of 16.03.2021)

CONTENTS

Introduction 4

1 Theory of probability. 5

2 Elements of mathematical statistics. 27

References. 38

Appendix A. Values of the Gauss function $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ 40

Appendix B. Values of the Laplace function $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ . . 42

Appendix C. Values of $t_\gamma = t(\gamma;\, n)$. 45

Appendix D. Critical points of chi-square distribution. 46

INTRODUCTION

One of the types of unaided work of students is the execution of calculation tasks during the semester. Their purpose is to develop practical skills. They are designed to help students more deeply capture theoretical material and learn how to apply the acquired knowledge to solve practical problems. The offered workshop with examples of calculation tasks corresponds to the course "Higher Mathematics" taught to students majoring in Power Engineering, Electrical Engineering and Electrical Mechanics full-time in the fourth module of the two-semester course. This workbook can also be used for students in other technical majors who study higher mathematics in a two-semester course.

This workbook covers virtually all major sections of the fourth module of the course on higher mathematics; it contains the necessary theoretical information, methodological instructions for completing the tasks, and examples of solving problems with a detailed explanation. A list of recommended literature is also provided. The manual also contains the necessary background material (tables, etc.). For ease of use, in this manual the numbering of examples corresponds to the numbering of tasks in the task book [14].

Calculation tasks are executed during the semester, which provides students with a systematic study of the course. In carrying out these tasks, students work with recommended textbooks and manuals, independently search for necessary literary sources and materials, analyze them and summarize, independently research and make written presentation of practical assignments.

The materials of this practicum can be considered as part of a synopsis of lectures on higher mathematics, in which the theoretical statements are illustrated in practical examples. That is why the materials in this manual can be used not only in the performance of calculation tasks, but also in preparation for the exam.

1 THEORY OF PROBABILITY

Events and their probabilities.

A phenomenon that occurs during an experiment is called an ***event***. Events are indicated by capital letters of the Latin alphabet: A, B, C ...

An event that inevitably occurs during an experiment is called a ***sure event*** (or ***certain event***). An event that can never happen during an experiment is called an ***impossible event*** (or ***null event***). An event that may or may not occur as a result of an experiment is called a ***random event***.

Two events are called ***mutually exclusive*** if the occurrence of one of them excludes the possibility of the occurrence of the other in the same experiment.

If the occurrence or non-occurrence of A does not affect the probability of occurrence of B, then A and B are said to be ***independent*** events.

Non-occurrence of the event A is called a ***complementary event*** (or failure of A) and it is denoted by $\overline{A}$.

Let n be the number of all possible outcomes of the experiment, and let m be the number of outcomes favorable to the event A. The ***probability*** of the event A in the classical sense is the ratio:

$$p(A) = \frac{m}{n}.$$

Accordingly, the probability of a certain event is 1, the probability of an impossible event is 0, the probability of a random event satisfies the double inequality $0 < p < 1$.

Let m outcomes (out of the total n outcomes) be favorable for the event A, then $n - m$ outcomes are favorable for the event $\overline{A}$. Then $p(A) + p(\overline{A}) = 1$.

Addition and multiplication of probabilities.

The ***sum*** (or the ***union***) of two events A and B is an event C in which at least one of the two events A and B occurs. The notation $C = A + B$ (or $C = A \cup B$) is used.

The ***product*** (or the ***intersection***) of two events A and B is an event C in which both events A and B occur simultaneously. The notation $C = AB$ (or $C = A \cap B$) is used.

If A and B are two events, the probability that A occurs given that B has occurred is denoted by $p_B(A)$ (or $p(A \mid B)$). This probability is called the ***conditional probability*** of A given that B has occurred. It is calculated by the formula:

$$p_B(A) = \frac{p(AB)}{p(B)}.$$

If two events A and B are mutually exclusive, then $p(AB) = 0$ and $p_B(A) = 0$, $p_A(B) = 0$. If two events A and B are independent, then $p_B(A) = p(A)$, $p_A(B) = p(B)$.

Addition theorem. The probability of the sum of two events is calculated by the formula:

$$p(A + B) = p(A) + p(B) - p(AB).$$

If A and B are mutually exclusive, then

$$p(A + B) = p(A) + p(B).$$

Multiplication theorem. The probability of the product of two events is calculated by the formula:

$$p(AB) = p(A) \cdot p_A(B) = p(B) \cdot p_B(A).$$

If A and B are independent, then

$$p(AB) = p(A) \cdot p(B).$$

Task 1.1. Solve the problem.

a) In a batch of 20 TVs, there are 5 defective TVs. To check 6 TVs are randomly selected. Find the probability that all selected TVs are working.

b) Two shooters make one shot at the target each. The probability that the first shooter will hit is 0.55. For the second shooter, this probability is equal to 0.6. What is the probability that at least one shooter will hit the target?

Solution. a) Number of ways to select 6 TVs from 20:

$$n = C_{20}^{6} = \frac{20!}{6!14!} = 38760.$$

This is the number of all possible outcomes.

Number of ways to select 6 working TVs from 15 working TVs:

$$m = C_{15}^{6} = \frac{15!}{6!9!} = 5005.$$

This is the number of outcomes favorable to the required event.

According to the classical definition of probability we have:

$$P = \frac{m}{n} = \frac{5005}{38760} = \frac{1001}{7752} \approx 0.1291.$$

b) Let the event A be that the first shooter hit the target. According to the condition of problem we have $P(A) = p_1 = 0.55$. Let the event B be that the second shooter hit the target. According to the condition of problem we have $P(B) = p_2 = 0.6$.

The miss of the first shooter (event $\overline{A}$) occurs with probability $q_1 = P(\overline{A}) = 1 - P(A) = 1 - p_1$. The miss of the second shooter (event $\overline{B}$) occurs with probability $q_2 = P(\overline{B}) = 1 - P(B) = 1 - p_2$.

Since the events $\overline{A}$, $\overline{B}$ are independent, then by theorem on the probability of the product of events we obtain: the probability of two simultaneous misses (event $\overline{A} \cdot \overline{B}$) is equal to

$$q = P(\overline{A} \cdot \overline{B}) = P(\overline{A}) \cdot P(\overline{B}) = q_1 q_2 = (1 - p_1)(1 - p_2).$$

Obviously, hitting a target with at least one shooter is the complementary event of the event $\overline{A} \cdot \overline{B}$. Then the required probability is equal to

$$p = 1 - q = 1 - (1 - p_1)(1 - p_2) = 1 - 0.45 \cdot 0.4 = 0.82.$$

Answer: a) $\frac{1001}{7752} \approx 0.1291$;

b) 0.82.

Total probability. Bayes' rule.

Let H_1, H_2, …, H_n be a set of mutually exclusive and exhaustive events, then this set is called a ***complete group of hypotheses***. For a complete group of hypotheses, the following equality holds: $p\left(\sum_{i=1}^{n} H_i\right) = 1$.

Theorem of total probability. Let event A occurs only if one hypothesis of a complete group H_1, H_2, …, H_n has occurred, then there is a formula of total probability:

$$p(A) = \sum_{i=1}^{n} p(H_i) p_{H_i}(A),$$

where $p(A)$ is a probability of the event A;

$p(H_i)$ is a probability of the hypothesis H_i;

$p_{H_i}(A)$ is a conditional probability of the event A given that H_i has occurred.

There is also ***Bayes' formula***:

$$p_A(H_k) = \frac{p(H_k) p_{H_k}(A)}{p(A)},$$

where $p_A(H_k)$ is the probability of realization of the hypothesis H_k given that A has already occurred.

Bayes' formula is also called the formula for reestimation of the probabilities of hypotheses ($p(H_k)$ is a ***probability a priori***, $p_A(H_k)$ is a ***probability a posteriori***).

Task 1.2. Solve the problem.

The driver could come to the gas station-1 with a probability of 45% or to the gas station-2 with a probability of 55%. If he fills the car at the gas station-1, he will be satisfied with the quality of fuel with a probability of 25%, and for the gas station-2 this probability is 95%. What is the probability that the driver stopped at a gas station-1, if it is known that he was satisfied with the quality of fuel?

Solution. The conditions of the problem correspond to the situation when Bayes' formula is used (a posteriori reestimations of hypotheses probabilities). Let H_1 be a hypothesis that the driver refueled the car at the gas station-1, and let H_2 be a hypothesis that the driver refueled the car at the gas station-2. According to the condition of the problem we have

$$p(H_1) = 0.45, \qquad p(H_2) = 0.55.$$

Let A be an event that means that the driver is satisfied with the quality of the fuel. The value of $p_{H_1}(A)$ is the conditional probability that the driver is satisfied with the quality of fuel, given that he refueled the car at the gas station-1; the value of $p_{H_2}(A)$ is the conditional probability that the driver is satisfied with the quality of fuel, given that he refueled the car at the gas station-2. According to the condition of the problem we have

$$p_{H_1}(A) = 0.25, \qquad p_{H_2}(A) = 0.95.$$

According to the formula of total probability we obtain

$$p(A) = p(H_1) \cdot p_{H_1}(A) + p(H_2) \cdot p_{H_2}(A) =$$
$$= 0.45 \cdot 0.25 + 0.55 \cdot 0.95 = 0.635.$$

The conditional probability $p_A(H_1)$ of hypothesis H_1, given that the event A has already occurred, according to Bayes' formula is equal to

$$p_A(H_1) = \frac{p(H_1) p_{H_1}(A)}{p(A)} = \frac{0.45 \cdot 0.25}{0.635} \approx 0.1772.$$

Note that it turned out that $p_A(H_1) \approx 0.1772 << p(H_1) = 0.45$. We reestimated the probability of the hypothesis to reduce it. In fact, *after it became known* that the driver is satisfied with the quality of fuel, the probability of arrival at gas station-1 should be reduced, because at this gas station the chances of being satisfied are much lower.

Answer: 0.1772 .

Sequence of independent trials (Bernoulli's scheme).

Bernoulli's experiment is one that possesses the following properties:

1) the experiment consists of n repeated trials;

2) each trial results is an outcome that may be classified as a ***success*** or a ***failure***;

3) the probability of a success, denoted by p, remains constant from trial to trial and repeated trials are independent.

Let the event A be a success in a single trial. Bernoulli's formula answers the question: "What is the probability that during n trials the event A will occur k $(0 \le k \le n)$ times?". ***Bernoulli's formula*** has the form:

$$P_n(k) = C_n^k p^k q^{n-k}, \qquad 0 \le k \le n,$$

where n is the number of trials;

p is the probability of success in a single trial;

q is the probability of failure in a single trial (i.e. $q = 1 - p$);

$C_n^k = \dfrac{n!}{k!(n-k)!}$ is the number of combinations of n elements by k (the number of ways to select k elements from those n elements that are available).

Note that $\sum_{k=0}^{n} P_n(k) = 1$.

Task 1.3. The probability that a student will pass the exam for the first time is 0.75. What is the probability that in a group of 8 students 7 students will pass the exam for the first time?

Solution. The conditions of the problem correspond to the Bernoulli's scheme (sequence of independent trials). We have $n = 8$,

$k = 7$, $p = 0.75$, $q = 1 - p = 0.25$. Then according to Bernoulli's formula we obtain:

$$P_8(7) = \frac{8!}{7!\cdot 1!}\cdot 0.75^7 \cdot 0.25^1 = 2\cdot 0.75^7 \approx 0.2670 .$$

Answer: 0.2670.

Asymptotic formulas for Bernoulli's scheme.

Bernoulli's formula gives the exact value of the probability of k occurrencies of the success during n trials. But for large values n, the use of this formula is difficult. Therefore, other formulas that give an approximate value of this probability are used.

If the number of trials n is large and the probability of a success in a single trial p is close to zero (success A is a ***rare event***), then the ***Poisson formula*** is used:

$$P_n(k) \approx \frac{\lambda^k}{k!} e^{-\lambda},$$

where $\lambda = np$. The formula is used when the following conditions are met: $n >> 1$, $p << 1$, $\lambda = np < 10$.

If the number of trials n is large and the probabilities p, q are neither close to zero nor to one, then ***Laplace's local theorem*** is used:

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{k - np}{\sqrt{npq}}\right),$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is ***Gauss function***. The function $\varphi(x)$ is tabulated, its values can be found in table (see Appendix A). When calculating, it should be borne in mind that the function is even: $\varphi(-x) = \varphi(x)$.

The generalization of the question in Bernoulli's scheme looks like this: "Find the probability that during n repeated trials the event A (a success in a single trial) will occur not less than k' times and not more than k'' times." The answer to this question is given by ***Laplace's integral theorem***:

$$P_n(k',k'') \approx \Phi\left(\frac{k''-np}{\sqrt{npq}}\right) - \Phi\left(\frac{k'-np}{\sqrt{npq}}\right),$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt$ is ***Laplace function***. The function $\Phi(x)$ is tabulated, its values can be found in table (see Appendix B). If $x > 5$, then $\Phi(x) = \Phi(+\infty) = \frac{1}{2}$ is assepted. When calculating, it should be borne in mind that the function is odd: $\Phi(-x) = -\Phi(x)$.

Task 1.4. During transportation, the detail can break with probability $p = 0.001$. $n = 3000$ details were transported. Find the probability that a) 4 details were broken; b) not more than 4 details were broken; c) more than 4 details were broken.

Solution. The conditions of the problem correspond to the Bernoulli's scheme (sequence of independent trials). We have: $n = 3000$, $p = 0.001$. But the use of Bernoulli's formula is difficult because the number n is very large. Besides, the value p is close to zero (the value q is close to one). However, the product $\lambda = np = 3000 \cdot 0.001 = 3 < 10$. Therefore, you can use the Poisson formula (the formula for the probability of rare events): $P_n(k) = \frac{\lambda^k}{k!}e^{-\lambda}$.

a) Probability that four details were broken ($k = 4$) is equal to:

$$P_n(4) = \frac{\lambda^4}{4!}e^{-\lambda} = \frac{\lambda^4}{24e^{\lambda}} = \frac{81}{24e^3} \approx 0.1680.$$

b) Events "no details were broken", "one detail was broken", etc. are mutually exclusive.

Probability that no details were broken ($k=0$ is number of broken ones) is equal to:

$$P_n(0)=\frac{\lambda^0}{0!}e^{-\lambda}=\frac{1}{e^{\lambda}}=\frac{1}{e^3}\approx 0.0498.$$

Probability that one detail was broken ($k=1$) is equal to:

$$P_n(1)=\frac{\lambda^1}{1!}e^{-\lambda}=\frac{\lambda}{e^{\lambda}}=\frac{3}{e^3}\approx 0.1494.$$

Probability that two details were broken ($k=2$) is equal to:

$$P_n(2)=\frac{\lambda^2}{2!}e^{-\lambda}=\frac{\lambda^2}{2e^{\lambda}}=\frac{9}{2e^3}\approx 0.2240.$$

Probability that three details were broken ($k=3$) is equal to:

$$P_n(3)=\frac{\lambda^3}{3!}e^{-\lambda}=\frac{\lambda^3}{6e^{\lambda}}=\frac{27}{6e^3}\approx 0.2240.$$

Probability that four details were broken ($k=4$) is equal to:

$$P_n(4)=\frac{\lambda^4}{4!}e^{-\lambda}=\frac{\lambda^4}{24e^{\lambda}}=\frac{81}{24e^3}\approx 0.1680.$$

Therefore, according to the theorem on the probability of the sum of mutually exclusive events, the probability that no more than 4 details were broken is equal to:

$$p=P_n(0)+P_n(1)+P_n(2)+P_n(3)+P_n(4)=$$

$$=0.0498+0.1494+0.2240+0.2240+0.1680=0.8153.$$

c) The probability of the complementary event (more than 4 details were broken) can be found as follows:

$$q=1-p=1-0.8153=0.1847.$$

Answer: a) 0.1680;
b) 0.8153;
c) 0.1847.

Task 1.5. a) During transportation, the detail can break with probability $p = 0.4$. $n = 3750$ details were transported. Find the probability that 1497 details were broken.

b) During transportation, the detail can break with probability $p = 0.5$. $n = 400$ details were transported. Find the probability that 1b) not less than 195 details and not more than 210 details were broken; 2b) not less than 195 details and not more than 390 details were broken.

Solution. a) The conditions of the problem correspond to the Bernoulli's scheme (sequence of independent trials). We have: $n = 3750$, $p = 0.4$, $q = 1 - p = 0.6$. But the use of Bernoulli's formula is difficult because the number n is very large. However, the product $\lambda = np = 3750 \cdot 0.4 = 1500 >> 10$. Therefore, Poisson's asymptotic cannot be used (this event is not a rare one). So, we apply the local Laplace theorem. We have:

$$\sqrt{npq} = \sqrt{3750 \cdot 0.4 \cdot 0.6} = 30,$$

$$\frac{k - np}{\sqrt{npq}} = \frac{1497 - 3750 \cdot 0.4}{30} = -0.1.$$

According to the table of values of the function $\varphi(x)$ (see Appendix A), taking into account its evenness, we obtain

$$\varphi\left(\frac{k - np}{\sqrt{npq}}\right) = \varphi(-0.1) = \varphi(0.1) = 0.3970.$$

Then by the local Laplace theorem we obtain

$$P_n(k) = \frac{1}{\sqrt{npq}} \cdot \varphi\left(\frac{k - np}{\sqrt{npq}}\right) = \frac{0.3970}{30} \approx 0.0132.$$

b) The conditions of the problem correspond to the Bernoulli's scheme (sequence of independent trials) with the parameters $n = 400$, $p = 0.5$, $q = 1 - p = 0.5$. But the use of Bernoulli's formula is difficult because the number n is very large. However, the product $\lambda = np = 400 \cdot 0.5 = 200 >> 10$. Therefore, Poisson's asymptotic cannot be used (this event is not a rare one). But the application of Laplace's local theorem is quite cumbersome, because when calculating the probability of the event "from 195 to 210 details were broken" it is necessary to add the probabilities of a large number of mutually exclusive events "195 details were broken", "196 details were broken", etc.:

$$P_{400}(195,210) = P_{400}(195) + P_{400}(196) + \cdots + P_{400}(210).$$

Therefore, we apply the Laplace integral theorem.

1b) We have:

$$\sqrt{npq} = \sqrt{400 \cdot 0.5 \cdot 0.5} = 10,$$

$$\frac{k' - np}{\sqrt{npq}} = \frac{195 - 400 \cdot 0.5}{10} = -0.5,$$

$$\frac{k'' - np}{\sqrt{npq}} = \frac{210 - 400 \cdot 0.5}{10} = 1.$$

According to the table of values of the function $\Phi(x)$ (see Appendix B), taking into account its evenness, we obtain

$$\Phi\left(\frac{k' - np}{\sqrt{npq}}\right) = \Phi(-0.5) = -\Phi(0.5) = -0.1915,$$

$$\Phi\left(\frac{k'' - np}{\sqrt{npq}}\right) = \Phi(1) = 0.3413.$$

Then, according to Laplace's integral theorem, the required probability is equal to

$$P_n(k', k'') = P_{400}(195,210) = \Phi(1) - \Phi(-0.5) =$$
$$= 0.3413 - (-0.1915) = 0.5328.$$

2b) Here we have the same:

$$\frac{k' - np}{\sqrt{npq}} = \frac{195 - 400 \cdot 0.5}{10} = -0.5,$$

$$\frac{k'' - np}{\sqrt{npq}} = \frac{390 - 400 \cdot 0.5}{10} = 19.$$

For $x = 19$ the value of $\Phi(x)$ is missing in the table. But if $x > 5$, then $\Phi(x) = \Phi(+\infty) = 0.5$ is assepted, then $\Phi(19) = 0.5$. According to Laplace's integral theorem, the required probability is equal to

$$P_n(k',k'') = P_{400}(195,390) = \Phi(19) - \Phi(-0.5) =$$
$$= 0.5 - (-0.1915) = 0.6915 .$$

Answer: a) $0,0132$;
b) 1b) 0.5328; 2b) 0.6915.

Random variables and their distributions

The term "statistical experiment" is used to describe any process by which several chance observations are obtained. All possible outcomes of an experiment comprise a set that is called the ***sample space***. We are interested in some numerical description of the outcome.

A ***random variable*** is a variable whose value is determined by the outcome of a random experiment. There are discrete and continuous random variables according to the set of its values.

A ***discrete random variable*** is one whose set of assumed values is countable (arises from *counting*).

A ***continuous random variable*** is one whose set of assumed values is uncountable (arises from *measurement*).

We shall use a capital X for the random variable and lowercase x_1, x_2, x_3... for the values of the random variable in an experiment. These x_i then represent an event that is a subset of the sample space. The probabilities of the events are given by $P(X = x_1) = p_1, P(X = x_2) = p_2, P(X = x_3) = p_3, ...$

A ***discrete probability distribution*** is a table (or a formula) listing all possible values that a discrete variable can take on, together with the associated probabilities. A piecewise linear graph constructed from this table is called a ***distribution polygon***. The following condition of normalization holds for the discrete distribution: $\sum_i p_i = 1$.

Analogous to the discrete probability distribution in the case of a continuous random variable is the ***probability density function***, which is equal to the probability dw of getting the value of x inside the interval dx per unit length of this interval: $f(x)=\frac{dw}{dx}$. Sometimes $f(x)$ is called the *differential distribution function.* Properties of probability density function $f(x)$:

1) $f(x)\geq 0$;

2) $P(\alpha < X < \beta)=\int_{\alpha}^{\beta} f(x)dx$;

3) $\int_{-\infty}^{\infty} f(x)dx=1$.

The last property is analogous to the discrete normalization condition $\sum_{i} p_i = 1$.

The function $F(x)=P(X<x)$ is called the ***distribution function*** (the *integral distribution function*) of a random variable. For a discrete random variable we have $F(x)=\sum_{(i)} p_i$; here the summation is performed by the numbers i for which $x_i < x$. For a continuous random variable we have $F(x)=\int dw = \int_{-\infty}^{x} f(\xi)d\xi$; here integration is performed for $\xi < x$. Notice that the distribution function for a continuous random variable is absolutely continuous over the whole real line. This is in contrast to the distribution function of a discrete random variable, which is a step function. However, whether discrete or continuous, any distribution function $F(x)$ must satisfy the following properties of a distribution function:

1) $F(-\infty)=0$, $F(+\infty)=1$;

2) $F(x)$ is a nondecreasing function;

3) $P(\alpha < X < \beta)=F(\beta)-F(\alpha)$.

Notice that $f(x)=F'(x)$ for a continuous random variable.

Numerical characteristics of a random variable.

Let X be a discrete random variable that takes values x_i with probabilities p_i respectively. Then the ***expected value*** of X denoted by $E(X)$, is defined as

$$E(X)=\sum_i x_i p_i .$$

To calculate this, we multiply each possible value of the variable by its probability, then add the results. The ***expected value*** of a continuous random variable X that has a probability density function $f(x)$ is given by

$$E(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx .$$

$E(X)$ is also called the ***mean*** of the probability distribution. Properties of an expected value $E(X)$:

1) $E(C)=C$, where $C=\text{const}$;

2) $E(CX)=CE(X)$, where $C=\text{const}$, X is a random variable;

3) $E(X+Y)=E(X)+E(Y)$, where X, Y are random variables;

4) $E(X\cdot Y)=E(X)\cdot E(Y)$, where X, Y are independent random variables.

Let X be a discrete random variable that takes values x_i with probabilities p_i respectively. Then the ***variance*** of X denoted by $V(X)$, is defined as

$$V(X)=\sum_i (x_i-E(X))^2 p_i .$$

Another way of calculating the variance is:

$$V(X)=\sum_i x_i^2 p_i-(E(X))^2 .$$

The ***variance*** of a continuous random variable X that has a probability density function $f(x)$ is given by

$$V(X)=\int_{-\infty}^{+\infty}(x-E(X))^2\cdot f(x)dx.$$

Another way of calculating the variance is:

$$V(X)=\int_{-\infty}^{+\infty}x^2\cdot f(x)dx-(E(X))^2.$$

Properties of a variance $V(X)$:

1) $V(C)=0$, where $C=\text{const}$;

2) $V(CX)=C^2\cdot V(X)$, where $C=\text{const}$, X is a random variable;

3) $V(X+Y)=V(X)+V(Y)$, where X, Y are independent random variables.

The ***standard deviation*** of the probability distribution is defined as

$$\sigma(X)=\sqrt{V(X)}.$$

The standard deviation is a number which describes the *spread* of the distribution. Small standard deviation means small spread, large standard deviation means large spread.

Task 1.6. Three cards were drawn at random from the deck (32 cards). Random variable is the number of extracted aces. Write the probability distribution for this random variable and find its expected value, variance and standard deviation.

Solution. Number of ways to choose three cards from the 32 cards available is $n=C_{32}^3$.

Calculate the probability that among the three drawn cards there aren't any aces. Since there are only 4 aces in the deck, there are 28 not aces. To avoid aces among the selected cards, you must choose them from these 28 cards. The number of ways to do this is equal to $m = C_{28}^3$. The required probability is:

$$p_0 = \frac{m}{n} = \frac{28!}{3!\cdot 25!} \cdot \frac{3!\cdot 29!}{32!} = \frac{26 \cdot 27 \cdot 28}{30 \cdot 31 \cdot 32} = \frac{13 \cdot 9 \cdot 7}{10 \cdot 31 \cdot 4} = \frac{819}{1240}.$$

Calculate the probability that among the three drawn cards there is one ace (accordingly, two cards are not aces). So, you need to choose one card out of 4 aces and regardless of that two cards out of 28. The required probability is:

$$p_1 = \frac{C_4^1 \cdot C_{28}^2}{C_{32}^3} = \frac{4!}{1!\cdot 3!} \cdot \frac{28!}{2!\cdot 26!} \cdot \frac{3!\cdot 29!}{32!} = \frac{3 \cdot 4 \cdot 27 \cdot 28}{30 \cdot 31 \cdot 32} = \frac{3 \cdot 9 \cdot 14}{10 \cdot 31 \cdot 4} = \frac{378}{1240}.$$

Calculate the probability that among the three drawn cards there are two aces (accordingly, one card is not ace). So, you need to choose two cards out of 4 aces and regardless of that one card more out of 28. The required probability is:

$$p_2 = \frac{C_4^2 \cdot C_{28}^1}{C_{32}^3} = \frac{4!}{2!\cdot 2!} \cdot \frac{28!}{1!\cdot 27!} \cdot \frac{3!\cdot 29!}{32!} = \frac{3 \cdot 2 \cdot 28 \cdot 2 \cdot 3}{30 \cdot 31 \cdot 32} = \frac{14 \cdot 3}{10 \cdot 31 \cdot 4} = \frac{42}{1240}.$$

Let's calculate the probability that all three drawn cards are aces. So, you need to choose three cards from 4 aces. The required probability:

$$p_3 = \frac{C_4^3}{C_{32}^3} = \frac{4!}{3!\cdot 1!} \cdot \frac{3!\cdot 29!}{32!} = \frac{4 \cdot 2 \cdot 3}{30 \cdot 31 \cdot 32} = \frac{1}{10 \cdot 31 \cdot 4} = \frac{1}{1240}.$$

Thus, the required probability distribution has the form:

x_i	0	1	2	3
p_i	$\frac{819}{1240}$	$\frac{378}{1240}$	$\frac{42}{1240}$	$\frac{1}{1240}$

Let's do a check: $p_0 + p_1 + p_2 + p_3 = \frac{819 + 378 + 42 + 1}{1240} = 1$.

Expected value:

$$E(X) = x_0 p_0 + x_1 p_1 + x_2 p_2 + x_3 p_3 =$$

$$= 0 \cdot \frac{819}{1240} + 1 \cdot \frac{378}{1240} + 2 \cdot \frac{42}{1240} + 3 \cdot \frac{1}{1240} = \frac{3}{8}.$$

Variance:

$$V(X) = x_0^2 p_0 + x_1^2 p_1 + x_2^2 p_2 + x_3^2 p_3 - (E(X))^2 =$$

$$= 0^2 \cdot \frac{819}{1240} + 1^2 \cdot \frac{378}{1240} + 2^2 \cdot \frac{42}{1240} + 3^2 \cdot \frac{1}{1240} - \left(\frac{3}{8}\right)^2 = \frac{609}{1984}.$$

Standard deviation:

$$\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{609}{1984}} \approx 0.5540.$$

Answer: probability distribution is the table; $E(X) = \frac{3}{8}$; $V(X) = \frac{609}{1984}$; $\sigma(X) \approx 0.5540$.

Task 1.7. A continuous random variable is given by a distribution function:

$$F(x) = \begin{cases} 0, & x < 0; \\ \frac{1}{16}x^2, & 0 \le x < 4; \\ 1, & x \ge 4. \end{cases}$$

Find the probability density function $f(x)$. Plot graphs of functions $F(x)$ and $f(x)$. Find the probability of the event $A = (X \in [a;b])$, where $a = \frac{1}{2}$, $b = 1$. For this random variable find its expected value, variance and standard deviation.

Solution. The probability density function is found by differentiating the distribution function $F(x)$ on each interval separately:

$$f(x)=F'(x)=\begin{cases}0, & x<0;\\ \frac{1}{8}x, & 0\le x<4;\\ 0, & x>4.\end{cases}$$

The graphs of these functions are schematically shown in Fig. 1.1.

The probability of event $A=(X\in[a;b])$ is found using the following formula (see property 3 of the distribution function):

$$P(X\in[a;b])=P(a<X<b)=F(b)-F(a),$$

$$P\left(X\in\left[\frac{1}{2};1\right]\right)=P\left(\frac{1}{2}<X<1\right)=F(1)-F\left(\frac{1}{2}\right)=$$

$$=\frac{1}{16}\cdot 1-\frac{1}{16}\cdot\left(\frac{1}{2}\right)^2=\frac{3}{64}=0.046875.$$

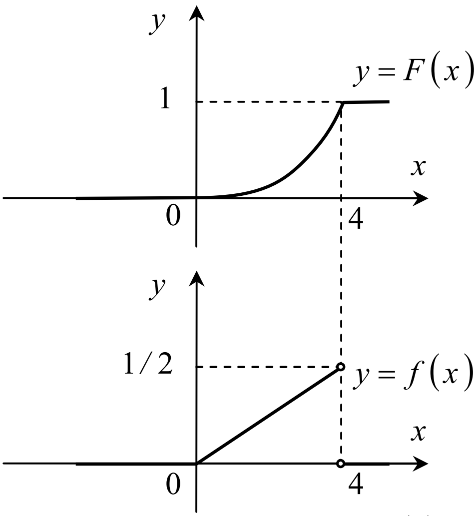

Figure 1.1 – The graphs of functions $F(x)$ and $f(x)$

Since the probability density function is given piecewise, using the additivity of the Riemann integral, we have:

$$E(X)=\int_{-\infty}^{+\infty} x\,f(x)dx=$$

$$=\int_{-\infty}^{0} x\,f(x)dx+\int_{0}^{4} x\,f(x)dx+\int_{4}^{+\infty} x\,f(x)dx=$$

$$=\underbrace{\int_{-\infty}^{0} x\cdot 0\,dx}_{=0}+\int_{0}^{4} x\,f(x)dx+\underbrace{\int_{4}^{+\infty} x\cdot 0\,dx}_{=0}=$$

$$=\int_{0}^{4} x\cdot\frac{1}{8}x\,dx=\left.\frac{x^3}{24}\right|_0^4=\frac{64}{24}-\frac{0}{24}=\frac{8}{3}.$$

Note that the value obtained coincides with the abscissa of the center of mass of the triangle in Fig. 1.1 (is two thirds of the length of its horizontal leg).

Similarly, the variance are calculated:

$$V(X)=\int_{-\infty}^{+\infty} x^2 f(x)dx-(E(X))^2=$$

$$=\int_{0}^{4} x^2\cdot\frac{1}{8}x\,dx-\left(\frac{8}{3}\right)^2=\left.\frac{x^4}{32}\right|_0^4-\frac{64}{9}=$$

$$=\frac{256}{32}-\frac{0}{32}-\frac{64}{9}=8-\frac{64}{9}=\frac{8}{9}.$$

Standard deviation:

$$\sigma(X)=\sqrt{V(X)}=\sqrt{\frac{8}{9}}=\frac{2\sqrt{2}}{3}\approx 0.9428.$$

Answer: $f(x)=\begin{cases}0, & x<0;\\ \frac{1}{8}x, & 0\le x<4;\\ 0, & x>4;\end{cases}$ $P\left(\frac{1}{2}<X<1\right)=\frac{3}{64}$;

$E(X)=\frac{8}{3}$; $V(X)=\frac{8}{9}$; $\sigma(X)=\frac{2\sqrt{2}}{3}\approx 0.9428$; graphs of functions $F(x)$ and $f(x)$ are shown in Fig. 1.1.

Normal probability distribution.

The normal probability distribution is very common in the field of statistics. Whenever you measure things like people's height, weight, salary, opinions or votes, the graph of the results obey a normal distribution very often.

A continuous random variable has a ***normal distribution*** with parameters a, σ, if its probability density function has the form

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

The distribution function has the form

$$F(x)=\frac{1}{2}+\Phi\left(\frac{x-a}{\sigma}\right),$$

where $\Phi(x)$ is Laplace function (see page 13).

The expected value and the variance of a continuous random variable normally distributed with the parameters a, σ, are equal to

$$E(X)=a, \qquad V(X)=\sigma^2.$$

It makes life a lot easier for us if we standardize the normal distribution, with an expected value of zero and a standard deviation of 1 unit. For the ***standardized normal distribution*** of $a=0$ and $\sigma=1$ we have:

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}=\varphi(x)$$ (it's Gauss function, see page 13);

$$F(x)=\frac{1}{2}+\Phi(x)$$ ($\Phi(x)$ is Laplace function, see page 13)

Thus, we can transform all the observations of any normal random variable X with mean a and variance σ to a new set of observations of another normal random variable Z with mean 0 and variance 1 using the following transformation: $Z=\dfrac{X-a}{\sigma}$.

The probability of hitting a continuous random variable X, normally distributed, in the interval from α to β is calculated by the formula:

$$P(\alpha < X < \beta)=\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right).$$

The probability that a normally distributed random variable X deviates from its mean in absolute value by no more than ε is calculated by the formula:

$$P(|X-a|<\varepsilon)=2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

Task 1.8. Solve the problem.

A large batch of beef entered the market. It is assumed that the weight of the carcasses is a random variable obeying the normal probability distribution with the mean $a=950$ kg and the standard deviation $\sigma=150$ kg. Determine the probability that the weight of a randomly selected carcass:

a) will be more than 1250 kg;
b) will be less than 850 kg;
c) will be between 800 and 1300 kg;
d) deviates from the mean by less than 50 kg;
e) deviates from the mean by more than 50 kg.

Solution. Let the random variable X be the weight of carcasses. Further, when solving, we use the properties of the function $\Phi(x)$ (see page 13) and the table of its values (see Appendix B).

a) $P(X > 1250) = P(1250 < X < +\infty) =$

$$= \Phi\left(\frac{+\infty - 950}{150}\right) - \Phi\left(\frac{1250 - 950}{150}\right) = \Phi(+\infty) - \Phi(2) =$$

$$= 0.5 - 0.4773 = 0.0227 .$$

b) $P(X < 850) = P(-\infty < X < 850) =$

$$= \Phi\left(\frac{850 - 950}{150}\right) - \Phi\left(\frac{-\infty - 950}{150}\right) = \Phi(-0.67) - \Phi(-\infty) =$$

$$= -\Phi(0.67) + \Phi(+\infty) = -0.2486 + 0.5 = 0.2514 .$$

c) $P(800 < X < 1300) = \Phi\left(\frac{1300 - 950}{150}\right) - \Phi\left(\frac{800 - 950}{150}\right) =$

$$= \Phi(2.33) - \Phi(-1) = \Phi(2.33) + \Phi(1) =$$

$$= 0.4901 + 0.3413 = 0.8314 .$$

d) $P(|X - 950| < 50) = 2\Phi\left(\frac{50}{150}\right) = 2\Phi(0.33) =$

$$= 2 \cdot 0.1293 = 0.2586 .$$

e) $P(|X - 950| > 50) = 1 - P(|X - 950| < 50) =$

$$= 1 - 0.2586 = 0.7414 .$$

Here we used the property of the probability of the complementary event.

Answer: a) 0.0227;

b) 0.2514;

c) 0.8314;

d) 0.2586;

e) 0.7414.

2 ELEMENTS OF MATHEMATICAL STATISTICS

Task 2.1. Random value X is an interest rate on deposits in city banks. A sample of 15 objects was made from the general population:

x: 1, 3, 2, 4, 8, 2, 7, 7, 4, 6, 4, 4, 6, 4, 3.

a) Form a variation series for this sample.
b) Form a statistical distribution.
c) Find the empirical distribution function and plot it.
d) Construct a frequency polygon.
e) Calculate the numerical characteristics of the sample: the sample mean, estimates for the median and mode, the sample variance, the unbiased standard deviation, the coefficient of variation. Explain the practical meaning of the values found.
f) On the basis of the frequency polygon and the calculated point estimates of the characteristics of the variation series make a conclusion about the law of distribution of the general population.

Solution.

a) The variation series (data values are ranked in increasing order):

1, 2, 2, 3, 3, 4, 4, 4, 4, 4, 6, 6, 7, 7, 8.

b) The below Table 2.1 is statistical distribution.

Here i is the ordinal number of the variant x_i.

The second column of the table contains the variants x_i.

n_i is the frequency of the corresponding variant x_i and it is determined from the original data.

Relative frequency ω_i is defined as the ratio of frequency n_i to sample size n:

$$\omega_i = \frac{n_i}{n} = \frac{n_i}{15}.$$

S_i is the accumulated relative frequency.

Table 2.1

i	x_i	n_i	$\omega_i = n_i / n$	S_i
1	1	1	1/15	1/15
2	2	2	2/15	3/15
3	3	2	2/15	5/15
4	4	5	5/15	10/15
5	6	2	2/15	12/15
6	7	2	2/15	14/15
7	8	1	1/15	1
		$\sum n_i = 15$	$\sum \omega = 1$	

c) The empirical distribution function has the form (we use the accumulated relative frequencies, Table 2.1):

$$F(x) = \begin{cases} 0, & if\ x \le 1 \\ 1/15, & if\ 1 < x \le 2 \\ 3/15 = 1/5, & if\ 2 < x \le 3 \\ 5/15 = 1/3, & if\ 3 < x \le 4 \\ 10/15 = 2/3, & if\ 4 < x \le 6 \\ 12/15 = 4/5, & if\ 6 < x \le 7 \\ 14/15, & if\ 7 < x \le 8 \\ 1, & if\ x > 8 \end{cases}$$

The graph of the empirical distribution function is shown in Fig. 2.1

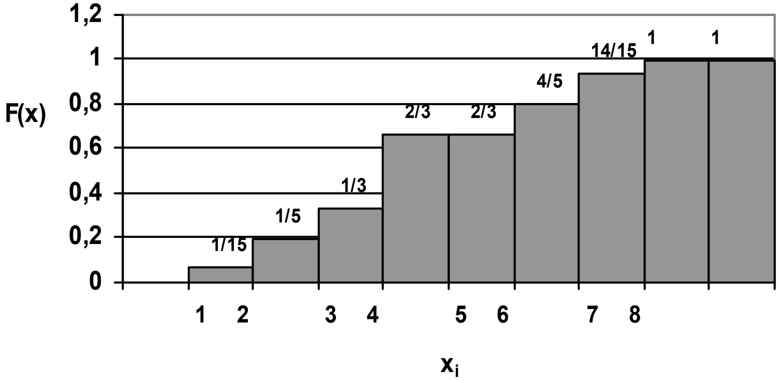

Figure 2.1 – The graphs of function $F(x)$

d) The frequency polygon is shown in Fig. 2.2.

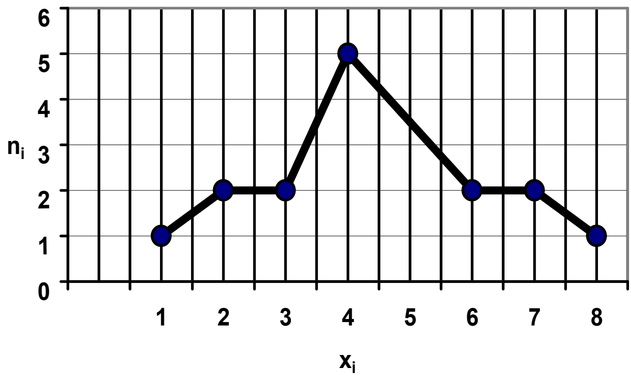

Figure 2.2 – The frequency polygon

e) Let's calculate the numerical characteristics of the sample.
The sample mean:

$$\overline{x} = \frac{1}{n}\sum x_i n_i = \frac{1}{15}(1 \cdot 1 + 2 \cdot 2 + \ldots + 8 \cdot 1) = 4.3333 .$$

The mean is a point estimate for the population expected value. This result means that the average interest rate on deposits in city banks is 4.3333%.

The sample variance:

$$V = \frac{1}{n}\sum x_i^2 n_i - \left(\overline{x}\right)^2 =$$

$$= \frac{1}{15}\left(1^2 \cdot 1 + 2^2 \cdot 2 + ... + 8^2 \cdot 1\right) - 4.3333^2 = 3.9556 .$$

The unbiased sample standard deviation:

$$s = \sqrt{\frac{n}{n-1}V} = \sqrt{\frac{15}{14} \cdot 3.9556} = 2.0587 .$$

The unbiased sample standard deviation is a point estimate for the population standard deviation. This result means that the average deviation of the interest rate on deposits in city banks from its average value is 2.0587%. That is, on average, the interest rate on deposits in city banks ranges from 2.2746% to 6.3920%.

Let's find the sample mode. This is the variant that corresponds to the maximum frequency (Table 2.1):

$$Mo = 4 .$$

It means that the most common (most frequent) value of the interest rate on deposits in city banks is 4%.

Let's find the sample median. This is the middle of the variation series:

$$Me = 4 .$$

It means that a half of the banks have an interest rate on deposits of less than 4%, and the other half of the banks have an interest rate on deposits of more than 4%.

The coefficient of variation:

$$\nu = \frac{\sqrt{V}}{\overline{x}} \cdot 100\% = \frac{\sqrt{3.9556}}{4.3333} \cdot 100\% = 45.90\% .$$

Since $\nu > 33\%$, the sample is not homogeneous (a high indicator of variability of X), the interest rate on deposits is influenced by various factors.

6) Based on the research data, it can be assumed that the general population is normally distributed.

Task 2.2. Let X be work experience (number of years) of employees of a certain company. It is known that X is normally distributed. A sample of $n = 9$ elements was made from the general population. The sample mean $\overline{x} = 17.4$ and the unbiased sample standard deviation $s = 3$ are found. Construct a 90% confidence interval for the unknown expected value if:

a) the population standard deviation $\sigma = 2.8$ is known;

b) the population standard deviation σ is unknown.

Solution.

a) The confidence interval for the expected value a is sought by the formula

$$\overline{x} - \delta < a < \overline{x} + \delta ,$$

where

$$\delta = t\frac{\sigma}{\sqrt{n}} .$$

Here t is the value of the argument of the Laplace function $\Phi(t)$ for which $\Phi(t) = \frac{\gamma}{2}$. We have:

$$\Phi(t) = \frac{0.9}{2} = 0.45 .$$

We find in the tables (see Appendix B): $t = 1.645$. The accuracy of the estimate:

$$\delta = 1.645\frac{2{,}8}{\sqrt{9}} = 1.5353 .$$

Then we get the confidence interval:

$$17.4 - 1.5353 < a < 17.4 + 1.5353 ,$$

$$15.8647 < a < 18.9353 .$$

б) The confidence interval for the expected value a is sought by the formula

$$\overline{x} - \delta < a < \overline{x} + \delta ,$$

where

$$\delta = t_\gamma \frac{s}{\sqrt{n}} .$$

We look for the critical point of the Student distribution $t_\gamma = t(n,\gamma)$ according to the tables (see Appendix C): $t_\gamma = t(n=9, \gamma = 0{,}9) = 1.86$. We have:

$$\delta = t_\gamma \frac{s}{\sqrt{n}} = 1.86 \cdot \frac{3}{\sqrt{9}} = 1.86 .$$

Then we get the confidence interval:

$$17.4 - 1.86 < a < 17.4 + 1.86 ,$$

$$15.54 < a < 19.26 .$$

Answer: a) $15.8647 < a < 18.9353$;

b) $15.54 < a < 19.26$.

Task 2.3. .The following data represent the mileage (km) of a taxi car with a full tank.

547, 547, 549, 540, 560, 551, 552, 553, 557, 549, 545, 546, 545, 548, 550, 551, 552, 555, 558, 548, 569, 542, 560, 548, 550, 552, 551, 557, 558, 551, 552, 557, 546, 549, 553, 548, 549, 550, 558, 551, 549, 543, 548, 535, 537, 550, 556, 555, 550, 548, 543, 546, 554, 533, 552, 556, 549, 546, 568, 554, 552, 551, 558, 559, 554, 557, 558, 552, 558, 526.

a) Make a grouped frequency table displaying class intervals, frequencies, relative frequencies, accumulated relative frequencies.

b) Draw a relative frequency histogram.

c) Draw a cumulative function.

d) Using the grouped data calculate the numerical characteristics of the sample: the sample mean, the sample variance, the unbiased standard deviation, the coefficient of variation.

e) Sustain or reject the null hypothesis of the normal distribution of the random variable using Pearson's χ^2 test with the significance level $\alpha = 0.05$.

Solution.

a) Size of the sample is $n = 70$. Let's form a grouped frequency table for this sample. The number of class intervals m is determined by Sturgess' rule:

$$m = 1 + 3.322\,lg\,n\,,$$
$$m = 1 + 3.32\,lg\,70 = 1 + 6.13 = 7.13\,.$$

Let's take the number of class intervals $m = 7$. Find the class width:

$$h = \frac{x_{max} - x_{min}}{m} = \frac{569 - 526}{7} \approx 7\,.$$

The below Table2.2 is a grouped frequency table.

Table 2.2

i	$h = 7$	x_i	n_i	$\omega_i = \frac{n_i}{n} = \frac{n_i}{70}$	S_i	$\frac{\omega_i}{h} = \frac{\omega_i}{7}$
1	523-530	526.5	1	0.0143	0.0143	0.0020
2	530-537	533.5	2	0.0286	0.0429	0.0041
3	537-544	540.5	5	0.0714	0.1143	0.0102
4	544-551	547.5	28	0.4	0.5143	0.0571
5	551-558	554.5	26	0.3714	0.8857	0.0531
6	558-565	561.5	6	0.0857	0.9714	0.0122
7	565-572	568.5	2	0.0286	1	0.0041
			$\sum n_i = 70$	$\sum \omega = 1$		

Here i is the ordinal number of the class interval.

The second column of the table contains the class intervals.

x_i is the midpoint of the corresponding class interval.

n_i is the frequency of the corresponding class interval and it is determined from the original data (the frequencies of values that fall on the border of two intervals are halved between these intervals).

Relative frequency ω_i is defined as the ratio of frequency n_i to sample size n:

$$\omega_i = \frac{n_i}{n} = \frac{n_i}{70}\,.$$

S_i is the accumulated relative frequency.

The last column of the table contains the data required to construct the histogram.

b) The relative frequency histogram is shown in Fig. 2.3. The heights of rectangles for each class $\frac{\omega_i}{h} = \frac{\omega_i}{7}$ are calculated in the last column of the Table 2.2.

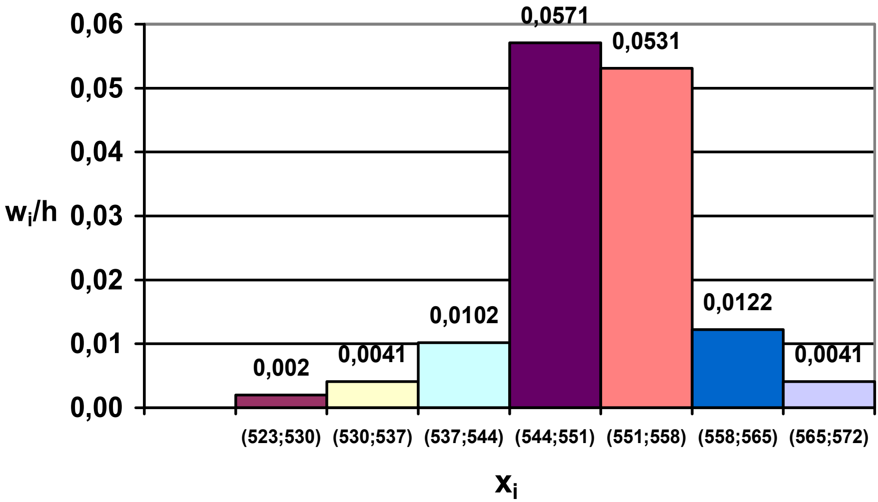

Figure 2.3 – The relative frequency histogram

c) The graph of the cumulative function $F(x)$ is shown in Fig. 2.4. We use the values of the accumulated relative frequencies S_i from the Table 2.2.

d) Let's calculate the numerical characteristics of the sample.

The sample mean:

$$\overline{x} = \frac{1}{n}\sum x_i n_i =$$

$$= \frac{1}{70}(526.5 \cdot 1 + 533.5 \cdot 2 + ... + 568.5 \cdot 2) = 550.7 .$$

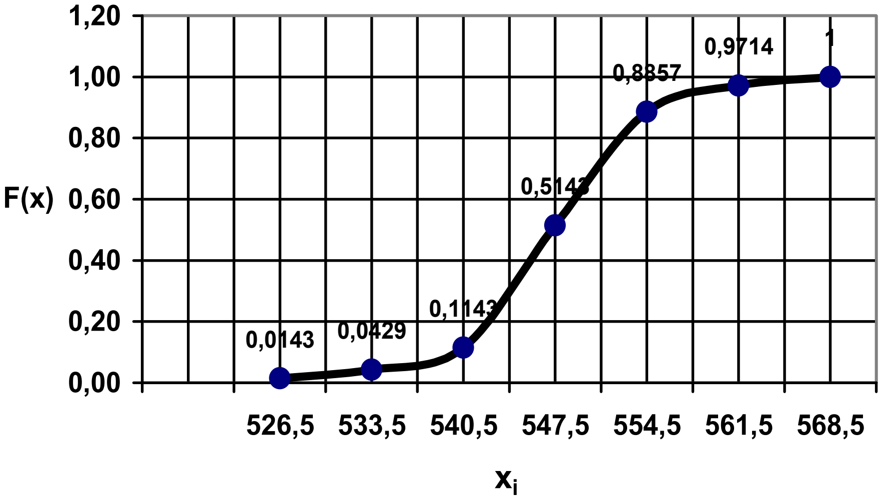

Figure 2.4 – The cumulative function

The sample variance:

$$V = \frac{1}{n}\sum x_i^2 n_i - \left(\overline{x}\right)^2 =$$

$$= \frac{1}{70}\left(526.5^2 \cdot 1 + 533.5^2 \cdot 2 + \ldots + 568.5^2 \cdot 1\right) - 550.7^2 = 52.76\,.$$

The sample standard deviation:

$$\hat{\sigma} = \sqrt{V} = \sqrt{52.76} = 7.2636\,.$$

The unbiased sample standard deviation:

$$s = \sqrt{\frac{n}{n-1}V} = \sqrt{\frac{70}{69} \cdot 52.76} = 7.3161\,.$$

The coefficient of variation:

$$\nu = \frac{\hat{\sigma}}{\overline{x}} \cdot 100\% = \frac{7.2636}{550.7} \cdot 100\% = 1.32\%\,.$$

e) Based on the research data, it can be assumed that the general population is normally distributed with parameters $a = \overline{x} = 550.7$, $\sigma = \hat{\sigma} = 7.2636$. We use Pearson's χ^2 test. This test of goodness of fit establishes whether an observed frequency distribution differs from a theoretical distribution. Let us calculate the theoretical frequencies of the normal distribution by the formula:

$$n_i' = \frac{nh}{\hat{\sigma}}\varphi(u_i) = \frac{70 \cdot 7}{7.2636}\varphi(u_i) = 67.4596\varphi(u_i).$$

Here $u_i = \dfrac{x_i - \overline{x}}{\hat{\sigma}}$; $\varphi(x)$ is Gauss function (see page 13), its values can be found in table (see Appendix A).

We make a calculation table:

Table 2.3

i	x_i	$u_i = \dfrac{x_i - \overline{x}}{\hat{\sigma}} = \dfrac{x_i - 550{,}7}{7{,}2636}$	$\varphi(u_i)$	$n_i' = 67{,}4596\varphi(u_i)$
1	526.5	-3.33	0.0016	0.1079
2	533.5	-2.37	0.0241	1.6258
3	540.5	-1.4	0.1497	10.0987
4	547.5	-0.44	0.3621	24.4271
5	554.5	0.52	0.3485	23.5097
6	561.5	1.49	0.1315	8.8709
7	568.5	2.45	0.0198	1.3357

Let's compare empirical and theoretical frequencies. We compose a calculation table from which we find the chi-squared test statistic:

$$\chi^2 = \sum_i \frac{(n_i - n_i')^2}{n_i'}.$$

Table 2.4

i	n_i		n'_i		$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	1	3	0.1079	1.7337	1.2663	1.6035	0.9249
2	2		1.6258				
3	5		10.0987		-5.0987	25.9968	2.5743
4	28		24.4271		3.5729	12.7655	0.5226
5	26		23.5097		2.4903	6.2018	0.2638
6	6		8.8709		-2.8709	8.2423	0.9291
7	2		1.3357		0.6643	0.4413	0.3304
$\sum$	100						$\chi^2 = 5.5452$

Here the frequencies of the first two intervals were united, since they are too small.

The number of class intervals is $s = 6$ (since two intervals were united). The number of parameters of the normal distribution determined from the sample is $r = 2$. Then the degrees of freedom are $k = s - r - 1 = 6 - 2 - 1 = 3$.

For the given significance level $\alpha = 0.05$ and the degrees of freedom $k = 3$, we find the threshold χ^2 value. This value can be found in table (see Appendix D):

$$\chi^2_{thr} = \chi^2(0.05;\ 3) = 7.81.$$

Since $\chi^2 < \chi^2_{thr}$, we do not reject the null hypothesis of the normal distribution of the general population. The discrepancy between empirical and theoretical frequencies is random. The null hypothesis is sustained with the significance level $\alpha = 0.05$, the observed data do not contradict the hypothesis of a normal distribution of the general population.

REFERENCES

1. Англо-русский словарь математических терминов / [под ред. П.С. Александрова]. – 2-е, исправл. и дополн. изд. – М. : Мир, 1994. – 416с.
2. Вентцель Е.С. Задачи и упражнения по теории вероятностей: учеб. пособие для втузов / Е.С. Вентцель, Л.А. Овчаров. – 5-е изд. исправл. – М.: Академия, 2003. – 448 с.
3. Вентцель Е.С. Теории вероятностей и её инженерные приложения: учеб. пособие для втузов / Е.С. Вентцель, Л.А. Овчаров. – 3-е изд. перераб. и доп. – М.: Академия, 2003. – 464 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб. пособие для вузов / В.Е. Гмурман. – Изд. 9-е, стер. – М.: Высш. шк., 2003. – 479 с.
5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учеб. пособие для студентов вузов / В.Е. Гмурман. – Изд. 5-е, стер. – М.: Высш. шк., 2001. – 400 с.
6. Гурский Е.И. Сборник задач по теории вероятностей и математической статистике / Е.И. Гурский. – Минск, “Вышэйш. школа”, 1975. –272 с.
7. Краткий англо-русский технический словарь / Ю.А. Кузьмин, В.А. Владимиров, Я.Л. Гельман и др. – Издательство: Московская международная школа переводчиков, 1992. – 416 с.
8. Сборник задач по по высшей математике с контрольными работами: 2 курс / К.Н. Лунгу, В.П. Норин, Д.Т.Письменный и др., под ред. С.Н. Федина. – М.: Айрис-пресс, 2007. – 592 с.
9. Микулик Н.А. Руководство к решению технических задач по теории вероятностей и математической статистике / Н.А. Микулик, Г.Н. Рейзина, под ред. Е.И. Гурского. – Минск, “Вышэйш. школа”, 1977 – 143 с.

10. Письменный Д.Т. Конспект лекций по терии вероятностей, математической статистике и случайным процессам / Д.Т. Письменный – М.: Айрис-пресс, 2008. – 288 с.
11. Bird J. Higher Engineering Mathematics / J. Bird. – AmsterdamTokyo : Newnes, 2010. – 705 p.
12. Kuznetsov B. Russian-English Polytechnic Dictionary / B. Kuznetsov. – Moscow: Russian Language Publishers, 1980. – 723 p.
13. Smirnov V.I. A course of Higher Mathematics / V.I. Smirnov; transl. from Russian by D.E. Brown. – Pergamon Press, 1964. – 543 p.
14. Calculation tasks on higher mathematics (4th module) for students majoring in 141 Power Engineering, Electrical Engineering and Electrical Mechanics / Compiler: Snizhko Nataliia. – Zaporizhzhia: Zaporizhzhia Polytechnic National University, 2021. – 37 p.

Appendix A

Values of the Gauss function $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

	0	1	2	3	4	5	6	7	8
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818

	0	1	2	3	4	5	6	7	8
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459
2.1	0440	0431	0422	0413	0404	0396	0387	0379	0371
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013
3,4	0012	0012	0012	0011	0011	0010	0014	0010	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003
3,8	0004	0004	0004	0004	0004	0004	0003	0003	0003
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001

Appendix B

Values of the Laplace function $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,19	0,0753	0,38	0,1480	0,57	0,2157
0,01	0,0040	0,20	0,0793	0,39	0,1517	0,58	0,2190
0,02	0,0080	0,21	0,0832	0,40	0,1554	0,59	0,2224
0,03	0,0120	0,22	0,0871	0,41	0,1591	0,60	0,2257
0,04	0,0160	0,23	0,0910	0,42	0,1628	0,61	0,2291
0,05	0,0199	0,24	0,0948	0,43	0,1664	0,62	0,2324
0,06	0,0239	0,25	0,0987	0,44	0,1700	0,63	0,2357
0,07	0,0279	0,26	0,1026	0,45	0,1736	0,64	0,2389
0,08	0,0319	0,27	0,1064	0,46	0,1772	0,65	0,2422
0,09	0,0359	0,28	0,1103	0,47	0,1808	0,66	0,2454
0,10	0,0398	0,29	0,1141	0,48	0,1844	0,67	0,2486
0,11	0,0438	0,30	0,1179	0,49	0,1879	0,68	0,2517
0,12	0,0478	0,31	0,1217	0,50	0,1915	0,69	0,2549
0,13	0,0517	0,32	0,1255	0,51	0,1950	0,70	0,2580
0,14	0,0557	0,33	0,1293	0,52	0,1985	0,71	0,2611
0,15	0,0596	0,34	0,1331	0,53	0,2019	0,72	0,2642
0,16	0,0636	0,35	0,1368	0,54	0,2054	0,73	0,2673
0,17	0,0675	0,36	0,1406	0,55	0,2088	0,74	0,2703
0,18	0,0714	0,37	0,1443	0,56	0,2123	0,75	0,2734

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,76	0,2764	0,99	0,3389	1,22	0,3883	1,45	0,4265
0,77	0,2794	1,00	0,3413	1,23	0,3907	1,46	0,4279
0,78	0,2823	1,01	0,3438	1,24	0,3925	1,47	0,4292
0,79	0,2852	1,02	0,3461	1,25	0,3944	1,48	0,4306
0,80	0,2881	1,03	0,3485	1,26	0.3952	1,49	0,4319
0,81	0,2910	1,04	0,3508	1,27	0.3980	1,50	0,4332
0,82	0.2939	1,05	0,3531	1,28	0,3997	1,51	0,4345
0,83	0,2967	1,06	0,3554	1,29	0,4015	1,52	0,4357
0,84	0,2995	1,07	0,3577	1,30	0,4032	1,53	0,4370
0,85	0,3023	1,08	0,3599	1,31	0,4049	1,54	0,4382
0,86	0,3051	1,09	0,3621	1,32	0,4066	1,55	0,4394
0,87	0,3078	1,10	0,3643	1,33	0,4082	1,56	0,4406
0,88	0,3106	1,11	0,3665	1,34	0,4099	1,57	0,4418
0,89	0,3133	1,12	0,3686	1,35	0,4115	1,58	0,4429
0,90	0,3159	1,13	0,3708	1,36	0,4131	1,59	0,4441
0,91	0,3186	1,14	0,3729	1,37	0,4147	1,60	0,4452
0,92	0,3212	1,15	0,3749	1,38	0,4162	1,61	0,4463
0,93	0,3238	1,16	0,3770	1,39	0,4177	1,62	0,4474
0,94	0,3264	1,17	0,3790	1,40	0,4192	1,63	0,4484
0,95	0,3289	1,18	0,3810	1,41	0,4207	1,64	0,4495
0,96	0,3315	1,19	0,3830	1,42	0,4222	1,65	0,4505
0,97	0,3340	1,20	0,3849	1,43	0,4230	1,66	0,4515
0,98	0,3365	1,21	0,3869	1,44	0,4251	1,67	0,4525

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,68	0,4535	1,91	0,4719	2,28	0,4887	2,74	0,4969
1,69	0,4545	1,92	0,4726	2,30	0,4893	2,76	0,4971
1,70	0,4554	1,93	0,4732	2,32	0,4898	2,78	0,4973
1,71	0,4564	1,94	0,4738	2,34	0,4904	2,80	0,4974
1,72	0,4573	1,95	0,4744	2,36	0,4909	2,82	0,4976
1,73	0,4582	1,96	0,475	2,38	0,4913	2,84	0,4977
1,74	0,4591	1,97	0,4756	2,40	0,4918	2,86	0,4979
1,75	0,4599	1,98	0,4761	2,42	0,4922	2,88	0,4980
1,76	0,4608	1,99	0,4767	2,44	0,4927	2,90	0,4981
1,77	0,4616	2,00	0,4772	2,46	0,4931	2,92	0,4982
1,78	0,4525	2,02	0,4783	2,48	0,4934	2,94	0,4984
1,79	0,4633	2,04	0,4793	2,50	0,4938	2,96	0,4985
1,80	0,4641	2,06	0,4803	2,52	0,4941	2,98	0,4985
1,81	0,4649	2,08	0,4812	2,54	0,4945	3,00	0,49865
1,82	0,4656	2,10	0,4821	2,56	0,4948	3,20	0,49931
1,83	0,4664	2,12	0,4830	2,58	0,4951	3,40	0,49966
1,84	0,4671	2,14	0,4838	2,60	0,4953	3,60	0,499841
1,85	0,4678	2,16	0,4846	2,62	0,4956	3,80	0,499928
1,86	0,4686	2,18	0,4854	2,64	0,4959	4,00	0,499968
1,87	0,4693	2,20	0,4861	2,66	0,4961	4,50	0,499997
1,88	0,4699	2,22	0,4868	2,68	0,4963	5,00	0,499997
1,89	0.4706	2,24	0,4875	2,70	0,4965		
1,90	0,4713	2,26	0,4881	2,72	0,4967		

Appendix C

Values of $t_\gamma = t(\gamma;\, n)$

n, sample size	γ, confidence level				
	0,90	**0,95**	**0,98**	**0,99**	**0,999**
6	2,02	2,57	3,36	4,03	6,87
7	1,94	2,45	3,14	3,71	5,96
8	1,89	2,36	3,00	3,50	5,41
9	1,86	2,31	2,90	3,36	5,04
10	1,83	2,26	2,82	3,25	4,78
11	1,81	2,23	2,76	3,17	4,59
12	1,80	2,20	2,72	3,11	4,44
13	1,78	2,18	2,68	3,05	4,32
14	1,77	2,16	2,65	3,01	4,22
15	1,76	2,14	2,62	2,98	4,14
16	1,75	2,13	2,60	2,95	4,07
17	1,75	2,12	2,58	2,92	4,02
18	1,74	2,11	2,57	2,90	3,97
19	1,73	2,10	2,55	2,88	3,92
20	1,73	2,09	2,54	2,86	3,88
21	1,72	2,09	2,53	2,85	3,85
22	1,72	2,08	2,52	2,83	3,82
23	1,72	2,07	2,51	2,82	3,79
24	1,71	2,07	2,50	2,81	3,77
25	1,71	2,06	2,49	2,80	3,75
26	1,71	2,06	2,49	2,79	3,73
27	1,71	2,06	2,49	2,78	3,71
28	1,70	2,05	2,47	2,77	3,69
29	1,70	2,05	2,47	2,76	3,67
30	1,70	2,05	2,46	2,76	3,66
31	1,70	2,04	2,46	2,75	3,65
41	1,68	2,02	2,42	2,70	3,55
51	1,68	2,01	2,40	2,68	3,50
61	1,67	2,00	2,39	2,66	3,46
121	1,66	1,98	2,36	2,62	3,37
∞	1,64	1,96	2,33	2,58	3,29

Appendix D
Critical points of chi-square distribution

k, number of degrees of freedom	α, significance level								
	0,99	**0,95**	**0,90**	**0,10**	**0,05**	**0,025**	**0,01**	**0,005**	**0,001**
1	0,0002	0,004	0,02	2,71	3,84	5,02	6,63	7,88	10,8
2	0,002	0,10	0,21	4,61	5,99	7,38	9,21	10,6	13,8
3	0,12	0,35	0,58	6,25	7,81	9,35	11,3	12,8	16,3
4	0,30	0,71	1,06	7,78	9,49	11,1	13,3	14,9	18,5
5	0,55	1,15	1,61	9,24	11,1	12,8	15,1	16,7	20,5
6	0,87	1,64	2,20	10,6	12,6	14,4	16,8	18,5	22,5
7	1,24	2,17	2,83	12,0	14,1	16,0	18,5	20,3	24,3
8	1,65	2,73	3,49	13,4	15,5	17,5	20,1	22,0	26,1
9	2,09	3,33	4,17	14,7	16,9	19,0	21,7	23,6	27,9
10	2,56	3,94	4,87	16,0	18,3	20,5	23,2	25,2	29,6
11	3,05	4,57	5,58	17,3	19,7	21,9	24,7	26,8	31,3
12	3,57	5,23	6,30	18,5	21,0	23,3	26,2	28,3	32,9
13	4,11	5,89	1,04	19,8	22,4	24,7	27,7	29,8	34,5
14	4,66	6,57	7,79	21,1	23,7	26,1	29,1	31,3	36,1
15	5,23	7,26	8,55	22,3	25,0	27,5	30,6	32,8	37,7
16	5,81	7,96	9,31	23,5	26,3	28,8	32,0	34,3	39,3
17	6,41	8,67	10,1	24,8	27,6	30,2	33,4	35,7	40,8
18	7,01	9,39	10,9	26,0	28,9	31,5	34,8	37,2	42,3
19	7,63	10,1	11,7	27,2	30,1	32,9	36,2	38,6	43,8
20	8,26	10,9	12,4	28,4	31,4	34,2	37,6	40,0	45,3
21	8,90	11,6	13,2	29,6	32,7	35,5	38,9	41,4	46,8
22	9,54	12,3	14,0	30,8	33,9	36,8	40,3	42,8	48,3
23	10,2	13,1	14,8	32,0	35,2	38,1	41,6	44,2	49,7
24	10,9	13,8	15,7	33,2	36,4	39,4	43,0	45,6	51,2
25	11,5	14,6	16,5	34,4	37,7	40,6	44,3	46,9	52,6
26	12,2	15,4	17,3	35,6	38,9	41,9	45,6	48,3	54,1
27	12,9	16,2	18,1	36,7	40,1	43,2	47,0	49,6	55,5
28	13,6	16,9	18,9	37,9	41,3	44,5	48,3	51,0	56,9
29	14,3	17,7	19,8	39,1	42,6	45,7	49,6	52,3	58,3
30	15,0	18,5	20,6	40,3	43,8	47,0	50,9	53,7	59,7